Stochastic Processes and Financial Mathematics (37363)

Chapter 4

Introduction to stochastic processes, Gaussian SPs, stationary SPs

Alex Novikov and Scott Alexander

School of Mathematical and Physical Sciences, UTS

Autumn 2025

Chapter outline

Topics:

- General definitions
- Gaussian processes
  - Definitions and basic examples
  - Brownian motion
  - Other examples
- Stationary processes

General definitions

The following diagram shows important classes of stochastic processes (SPs) and some important connections between them.

Gaussian | Stationary | Markov | Renewal .... Martingales ....

Gaussian-Markov → Gaussian-Markov-Stationary,Ornstein-Uhlenbeck

Compound Poisson processes → Levy processes

Continuous Markov (Diffusion processes) | Jump-diffusion

Time Series

General definitions

Definition 1 (stochastic process)

A stochastic process (SP) $X(t, \omega)$, $t \in D$ and $\omega \in \Omega$, is a collection of measurable RVs indexed by time t on a probability space (Ω, F, P), where the set D is called the index set of the process.

So a stochastic process is a function of both time t and elementary event ω, but for the most part the dependence on ω can be left implicit and we write alternatively

$$X_t \equiv X(t) \equiv X(t, \omega), \quad t \in D.$$

Other notation is frequently employed to define a SP, such as

$$X = (X_t)_{t \in D}.$$

General definitions

A SP X_t, $t \in D$, can be considered a mapping

$$X : D \times \Omega \to R$$

where $R \subseteq \mathbb{R}$, i.e. some subset of the real numbers $\mathbb{R}$.

The SP X_t is called

- **discrete-time** if D is countable (e.g. the set of integers $\mathbb{Z}$, set of natural numbers $\mathbb{N}$, some countable subset of $\mathbb{R}$ etc.)
- **continuous-time** if D is uncountable (e.g. the set of real numbers $\mathbb{R}$, non-negative real numbers $\mathbb{R}_{\geq 0}$ or some interval of $\mathbb{R}$ etc.)
- **discrete-space** if R is countable
- **continuous-space** if R is uncountable.

Vector SPs can also be defined.

General definitions

If $t = t^* \in D$ is fixed then $X(t^*, \omega)$ is a RV.

If $\omega = \omega^* \in \Omega$ is fixed then $X(t, \omega^*)$ is a trajectory (or realisation or path).

Definition 2 (finite-dimensional distributions FDD)

A family of finite-dimensional distributions on (Ω, F, P) is a function

$$F_{X_{t_1},\dots,X_{t_n}}(x_1, \dots x_n) = P(X_{t_1} \le x_1, \dots, X_{t_n} \le x_n),$$

$t_i \in D$ and $x_i \in \mathbb{R}$, consistent in the sense that

$$\lim_{x_1 \to \infty} F_{X_{t_1},\dots,X_{t_n}}(x_1, \dots x_n) = F_{X_{t_2},\dots,X_{t_n}}(x_2, \dots x_n),$$
$$\lim_{x_n \to \infty} F_{X_{t_1},\dots,X_{t_n}}(x_1, \dots x_n) = F_{X_{t_1},\dots,X_{t_{n-1}}}(x_1, \dots x_{n-1}).$$

Question.
From the definition above we see that given a probability space (Ω, F, P) we can define a stochastic process X_t with family of FDDs $F_{X_{t_1},\dots,X_{t_n}}(x_1, \dots x_n)$.

What about the reverse though? Given a family of finite FDDs $F_{X_{t_1},\dots,X_{t_n}}(x_1, \dots x_n)$, can we construct a probability space (Ω, F, P) and stochastic process X_t consistent with this family of FDDs?

The answer is "yes", via the Kolmogorov and Caratheodory extension theorems – see [Chung, 2001].

Why is this important?

We can define the behaviour of a SP via a family of FDDs and from this construct a SP with these properties that is supported by all the machinery of a probability space.

General definitions

Before looking at some examples we introduce some notation.

Notation.
Set the mean function

$$m(t) = E[X_t]$$

and covariance function

$$\begin{aligned} Q(t,s) &= \mathrm{cov}(X_t, X_s) \\ &= E\big[(X_t - E[X_t])(X_s - E[X_s])\big] \\ &= E[X_t X_s] - E[X_t]E[X_s]. \end{aligned}$$

Gaussian processes – definitions and basic examples

Definition 3 (Gaussian process)

A stochastic process X_t, $t \in D$, is called a Gaussian process if any random vector $\boldsymbol{X} = (X_{t_1}, \dots, X_{t_n})^T$ is joint Gaussian.

Set

$$\boldsymbol{m} = E[\boldsymbol{X}] := \big(E[X_{t_1}], \dots, E[X_{t_n}]\big)^T = \begin{pmatrix} m(t_1) \\ \vdots \\ m(t_n) \end{pmatrix}$$

and

$$\begin{aligned} \boldsymbol{Q} &= \operatorname{cov}(\boldsymbol{X}, \boldsymbol{X}) = E\big[(\boldsymbol{X} - E[\boldsymbol{X}])(\boldsymbol{X} - E[\boldsymbol{X}])^T\big] \\ &= \big(\operatorname{cov}(X_{t_i}, X_{t_j})\big)_{1 \le i,j \le n} = \begin{pmatrix} Q(t_1, t_1) & \cdots & Q(t_1, t_n) \\ \vdots & \ddots & \vdots \\ Q(t_n, t_1) & \cdots & Q(t_n, t_n) \end{pmatrix}. \end{aligned}$$

Gaussian processes – definitions and basic examples

It follows from Chapter 2 that such a Gaussian SP has CF

$$E[e^{i\boldsymbol{u}\cdot\boldsymbol{X}}] = E\left[e^{i\sum_{k=1}^{n} u_k X_{t_k}}\right] = e^{i\boldsymbol{m}\cdot\boldsymbol{u} - \frac{1}{2}\boldsymbol{u}\cdot\boldsymbol{Q}\boldsymbol{u}}$$

$$= \exp\left(i\sum_{k=1}^{n} u_k m(t_k) - \frac{1}{2}\sum_{k=1}^{n}\sum_{j=1}^{n} u_k u_j Q(t_k, t_j)\right).$$

The following proposition follows.

Proposition 1 (Gaussian SP distribution)

If X_t, $t \in D$, is a Gaussian process then to determine any joint distribution

$$F_{X_{t_1},\dots,X_{t_n}}(x_1, \dots, x_n) = P(X_{t_1} \le x_1, \dots, X_{t_n} \le x_n)$$

it is sufficient to know only two functions,

$$m(t) = E[X_t] \quad \text{and} \quad Q(t,s) = \mathrm{cov}(X_t, X_s).$$

We also know from Chapter 2 that if $\boldsymbol{u} \cdot \boldsymbol{Q}\boldsymbol{u} > 0$ for any $\boldsymbol{u} \neq \boldsymbol{0}$ then there exists the joint-PDF

$$\begin{aligned} f_{\boldsymbol{X}}(\boldsymbol{u}) &\equiv f_{X_{t_1},\ldots,X_{t_n}}(u_1,\ldots,u_n) \\ &= \frac{1}{\sqrt{(2\pi)^n \det(\boldsymbol{Q})}} e^{-\frac{1}{2}(\boldsymbol{u}-\boldsymbol{m})\cdot\boldsymbol{Q}^{-1}(\boldsymbol{u}-\boldsymbol{m})}. \end{aligned}$$

For example, if $\mathrm{var}(X_t) = Q(t,t) > 0$ then

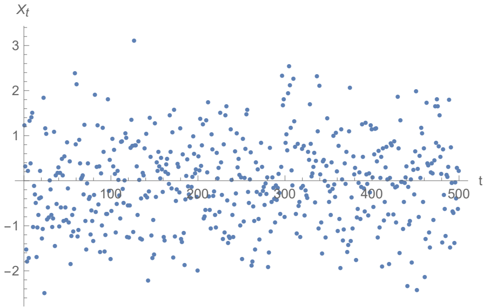
$$f_{X_t}(x) = \frac{1}{\sqrt{2\pi Q(t,t)}} e^{-\frac{(x-m(t))^2}{2Q(t,t)}}.$$

Example (discrete time Gaussian white noise).
Let $D = \{0, \Delta, 2\Delta, \ldots\}$, $\Delta > 0$ and $X_t \sim N(m, \sigma^2)$ with the X_t independent RVs. In this case

$$m(t) = E[X_t] = m \quad \text{and} \quad Q(t,s) = \begin{cases} \sigma^2 & \text{if } t = s \\ 0 & \text{if } t \neq s \end{cases}.$$

Below is a Mathematica example for the case $m = 0$, $\sigma^2 = 1$, $\Delta = 1$ and $T = 500$.

```
SeedRandom[123];
TT = 500;
data = Table[RandomReal[NormalDistribution[]], {TT}];
ListPlot[data, PlotMarkers → {Automatic, 5}, AxesLabel → {"t", "Xt"}]
```



Example (discrete time Gaussian random walk).
Let X_t be a Gaussian white noise process. Then the process

$$Y_t = X_1 + \cdots + X_t = \sum_{k=1}^{t} X_k, \quad Y_0 = 0,$$

is a discrete time Gaussian random walk.

To describe the distribution of this process we need

$$E[Y_t] = E\Big[\sum_{k=1}^{t} X_k\Big] = \sum_{k=1}^{t} E[X_k] = mt$$

and

$$\mathrm{var}(Y_t) = \mathrm{var}\left(\sum_{k=1}^{t} X_k\right) = \sum_{k=1}^{t} \mathrm{var}(X_k) = \sigma^2 t$$

so $Y_t \sim N(mt, \sigma^2 t)$.

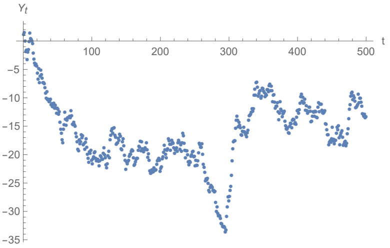
Gaussian processes – definitions and basic examples

It can also be shown that the covariance

$$Q(t, s) = \mathrm{cov}(Y_t, Y_s) = \sigma^2 \min(t, s).$$

Below is a Mathematica example for the case $m = 0$, $\sigma^2 = 1$, $\Delta = 1$ and $T = 500$.

```
SeedRandom[123];
TT = 500;
data = Table[RandomReal[NormalDistribution[]], {TT}];
data = Accumulate[data];
ListPlot[data, PlotMarkers → {Automatic, 5}, AxesLabel → {"t", "Y_t"}]
```



Gaussian processes – Brownian motion

A central object of study in this subject is a fundamentally important SP called Brownian motion.

Definition 4 (Brownian motion)

A stochastic process W_t, $t \in [0, \infty)$, is said to be a Brownian motion (BM), or Wiener process, if W_t is a Gaussian process with

$$E[W_t] = mt \quad \text{and} \quad Q(t,s) = \sigma^2 \min(t,s).$$

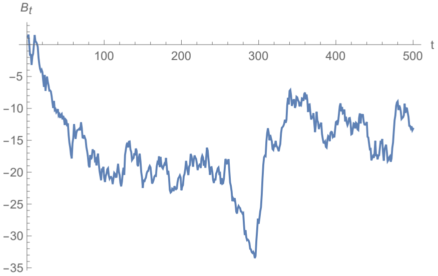
That is, $W_t \sim N(mt, \sigma^2 t)$.

When $m = 0$ and $\sigma^2 = 1$, the process is called a standard Brownian motion, which we denote B_t.

There are many analytical tools that have been developed in order to work with BM.

Example (Brownian motion).
Below is a Mathematica example for a standard BM with $T = 500$.

```
SeedRandom[123];
TT = 500;
data = Table[RandomReal[NormalDistribution[]], {TT}];
data = Accumulate[data];
ListPlot[data, Joined → True, AxesLabel → {"t", "Bt"}]
```



Note that the best we can do with a PC is a discrete-time approximation (in this case with linear interpolation between points).

Gaussian processes – Brownian motion

It's all well and good to define mathematical objects, but can we actually construct one?

Proposition 2 (existence of Brownian motion)

Let $\xi_k \sim N(0,1)$, $k = 0, 1, 2, ...$, be independent RVs. Then the process

$$X_t = \frac{\xi_0 t}{\sqrt{\pi}} + \sum_{k=1}^{\infty} \xi_k \frac{2\sin(kt)}{k\sqrt{\pi}}$$

is a standard Brownian motion for $t \in [0, \pi]$.

The proof involves use of Fourier series, which is beyond the scope of the subject.

Gaussian processes – Brownian motion

The first property we establish is the distribution of increments of BM.

Proposition 3 (distribution of increments of BM)

For $\Delta > 0$ the increment $(W_{t+\Delta} - W_t) \sim N(m\Delta, \sigma^2\Delta)$.

Proof.
We know that W_t is Gaussian and by Definition 4

$$\begin{aligned} E[W_{t+\Delta} - W_t] &= E[W_{t+\Delta}] - E[W_t] = m(t+\Delta) - mt \\ &= m\Delta \end{aligned}$$

with

$$\begin{aligned} \mathrm{var}(W_{t+\Delta} - W_t) &= \mathrm{var}(W_{t+\Delta}) + \mathrm{var}(W_t) - 2\,\mathrm{cov}(W_{t+\Delta}, W_t) \\ &= \sigma^2(t+\Delta) + \sigma^2 t - 2\sigma^2 t = \sigma^2\Delta \end{aligned}$$

so $(W_{t+\Delta} - W_t) \sim N(m\Delta, \sigma^2\Delta)$.

Gaussian processes – Brownian motion

The next property we establish is the independence of non-overlapping increments.

Proposition 4 (independence of increments of BM)

For any disjoint intervals $[t, t + \Delta_1]$ and $[s, s + \Delta_2]$

$$\operatorname{cov}(W_{t+\Delta_1} - W_t, W_{s+\Delta_2} - W_s) = 0$$

implying their independence.

Proof.

Consider the case $t + \Delta_1 > t > s + \Delta_2 > s$. Then

$$\begin{aligned}\operatorname{cov}(W_{t+\Delta_1} - W_t, W_{s+\Delta_2} - W_s) &= \operatorname{cov}(W_{t+\Delta_1}, W_{s+\Delta_2}) \\ &\quad - \operatorname{cov}(W_{t+\Delta_1}, W_s) - \operatorname{cov}(W_t, W_{s+\Delta_2}) + \operatorname{cov}(W_t, W_s)\end{aligned}$$

Gaussian processes – Brownian motion

$$
\begin{aligned}
&= \sigma^2\big(\min(t+\Delta_1, s+\Delta_2) - \min(t+\Delta_1, s) - \min(t, s+\Delta_2)\big) \\
&\quad + \min(t,s)\big) \\
&= \sigma^2\big((s+\Delta_2) - s - (s+\Delta_2) + s\big) = 0.
\end{aligned}
$$

As $W_{t+\Delta_1} - W_t$ and $W_{s+\Delta_1} - W_s$ have a joint Gaussian distribution, zero covariance implies their independence (see TNC Chapter 2).

Now we look at how BM scales.

Proposition 5 (self-similarity or scaling of BM)

If B_t is a standard BM then the process

$$X_t = \sqrt{a} B_{t/a}, \quad t \geq 0,$$

is also a standard Brownian motion for any constant $a > 0$.

Proof.
We see that

$$E[X_t] = E[\sqrt{a} B_{t/a}] = \sqrt{a} E[B_{t/a}] = 0$$

and

$$\begin{aligned} \operatorname{cov}[X_t, X_s] &= \operatorname{cov}(\sqrt{a} B_{t/a}, \sqrt{a} B_{s/a}) = a \operatorname{cov}(B_{t/a}, B_{s/a}) \\ &= a \min\left(\frac{t}{a}, \frac{s}{a}\right) = \min(t, s) \end{aligned}$$

so X_t is a standard BM by Definition 4.

Gaussian processes – Brownian motion

For the next property we require an interim result, a theorem that we state without proof.

Theorem 1 (Kolmogorov's criterion)

Let X_t, $t \in [0, T]$, be a real-valued process. If there exists an $\alpha > 0$ and $\varepsilon > 0$ so that for any $0 \leq u \leq t \leq T$

$$E[|X_t - X_u|^\alpha] \leq c(t-u)^{1+\varepsilon}$$

for some constant $c > 0$, then there exists a version of X_t with continuous sample paths (i.e. which are Hölder continuous of order $h < \varepsilon/\alpha$).

Now we can show that BM has continuous trajectories.

Gaussian processes – Brownian motion

Proposition 6 (continuity of trajectories of BM)

Brownian motion has a version with continuous (almost surely) trajectories.

Proof.
We demonstrate this property for standard BM (it generalises to the general case without too much difficulty).

Note for any $u < t$, $B_t - B_u \sim N(0, t-u)$ and therefore

$$E[|B_t - B_u|^4] = 3(t-u)^2.$$

So we can set $\alpha = 4$, $c = 3$ and $\varepsilon = 1$ in Kolmogorov's criterion and claim that B_t is Hölder continuous of order $h < 1/4$.

So we have continuity (this is often assumed by definition), but next we see that BM is not differentiable.

Gaussian processes – Brownian motion

Proposition 7 (non-differentiability of trajectories of BM)

Brownian motion is not differentiable at any point.

Proof.
We demonstrate this property for standard BM (it generalises to the general case without difficulty).

First note that by Proposition 3, $(B_{t+\Delta} - B_t) \sim N(0, \Delta)$ so

$$\frac{B_{t+\Delta} - B_t}{\Delta} \sim N\Big(0, \frac{1}{\Delta}\Big).$$

By results in Chapter 2 this means

$$X = \sqrt{\Delta}\frac{B_{t+\Delta} - B_t}{\Delta} \sim N(0, 1).$$

Gaussian processes – Brownian motion

So for any $\varepsilon > 0$

$$
\begin{aligned}
P\Big(\Big|\frac{B_{t+\Delta} - B_t}{\Delta}\Big| > \varepsilon\Big) &= P\Big(\Big|\sqrt{\Delta}\frac{B_{t+\Delta} - B_t}{\Delta}\Big| > \sqrt{\Delta}\varepsilon\Big) \\
&= P\big(|X| > \sqrt{\Delta}\varepsilon\big) = P\big(\{X < -\sqrt{\Delta}\varepsilon\} \cup \{X > \sqrt{\Delta}\varepsilon\}\big) \\
&= \int_{-\infty}^{-\sqrt{\Delta}\varepsilon} f_X(x)dx + \int_{\sqrt{\Delta}\varepsilon}^{\infty} f_X(x)dx \\
&\to \int_{-\infty}^{\infty} f_X(x)dx = 1
\end{aligned}
$$

as $\Delta \to 0$.

So $(B_{t+\Delta} - B_t)/\Delta$ does not have a limit (in probability).

Gaussian processes – other examples

A process that is gaining popularity in mathematical finance applications is fractional Brownian motion.

Definition 5 (fractional Brownian motion)

A stochastic process B_t^H, $t \in [0, \infty)$, is said to be a fractional Brownian motion (fBm) with Hurst exponent $H \in (0, 1]$ if B_t^H is a Gaussian process with

$$E[B_t^H] = 0 \quad \text{and} \quad E\left[|B_t^H - B_s^H|^2\right] = |t - s|^{2H}$$

and $B_t^H \sim N(0, t^{2H})$.

As

$$E\left[|B_t^H - B_s^H|^2\right] = E\left[(B_t^H)^2\right] + E\left[(B_s^H)^2\right] - 2E\left[B_t^H B_s^H\right]$$

we get

$$Q(t, s) = E\left[B_t^H B_s^H\right] = \frac{1}{2}(t^{2H} + s^{2H} - |t - s|^{2H}).$$

Gaussian processes – other examples

If

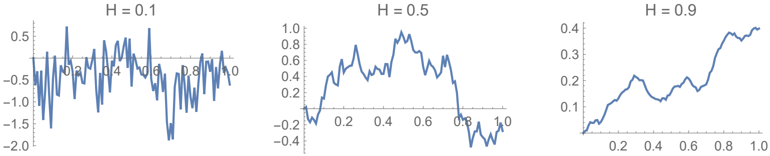
- $H \in (0, \frac{1}{2})$ then increments have negative correlation
- $H = \frac{1}{2}$ then $B_t^{1/2}$ is a standard BM
- $H \in (\frac{1}{2}, 1)$ then increments have positive correlation.
- $H = 1$ then $B_t^1 = \xi t$ where $\xi \sim N(0,1)$

To simulate continuous time Gaussian processes one may use

- discretisation in time with use of (e.g.) Cholesky decomposition;
- truncated Karhounen-Loev expansions
- inbuilt Mathematica functions!

Example (Fractional Brownian motion).
Below is a Mathematica example for fBMs.

```
TT = 1;
μ = 0;  σ = 1;
fbm := FractionalBrownianMotionProcess[μ, σ, h];
Table[ListPlot[RandomFunction[fbm, {0, 1, 0.01}],
   Joined → True, PlotLabel → "H = " <> ToString[h]], {h, {0.1, 0.5, 0.9}}
 ] // GraphicsRow
```



Gaussian processes – other examples

Another process used in mathematical finance applications is the Brownian bridge.

Definition 6 (Brownian bridge)

A stochastic process B_t^b, $t \in [0,1]$, is said to be a (standard) Brownian bridge if it is a Gaussian process with

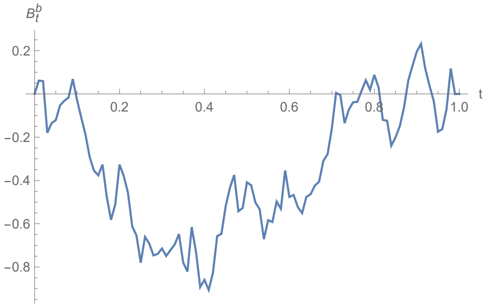
$$E[B_t^b] = 0 \quad \text{and} \quad Q(t,s) = \mathrm{cov}(B_t^b, B_s^b) = \min(t,s) - ts,$$

i.e. $B_t^b \sim N(0, t - t^2)$.

Note that as $\mathrm{var}(B_0^b) = \mathrm{var}(B_1^b) = 0$ we may assume that $B_0^b = B_1^b = 0$.

Example (Standard Brownian bridge).
Below is a Mathematica example for a standard Brownian bridge.

```
TT = 1;
data = RandomFunction[BrownianBridgeProcess[], {0, TT, .01}];
ListLinePlot[data, AxesLabel → {"t", "B_t^b"}]
```



Gaussian processes – other examples

It is possible to use standard BMs to construct a Brownian bridge, as the next exercise demonstrates.

Exercise.
Let $B_t \equiv B(t)$ be a standard BM. Show that processes

$$X_t^{(1)} = B_t - tB_1, \quad t \in [0,1],$$

and

$$X_t^{(2)} = (1-t)B\left(\frac{t}{1-t}\right), \quad t \in [0,1),$$

are Brownian Bridges.

Solution. As class work.

Gaussian processes – other examples

A common process used in econometrics is the autoregressive process.

Definition 7 (Gaussian autoregressive process)

An autoregressive process of order 1 (AR(1) process) is a solution of

$$X_t = \lambda X_{t-1} + e_t, \quad t \in \{1, 2, \dots\}$$

where $e_t \sim N(0, \sigma^2)$ and independent RVs and X_0 may also be a Gaussian RV independent of the e_t.

An autoregressive process of order p (AR(p) process) is a solution of

$$X_t = \lambda_1 X_{t-1} + \lambda_2 X_{t-2} + \cdots \lambda_p X_{t-p} + e_t, \quad t \in \{p, p+1, \dots\}.$$

Note by back-substitution the AR(1) process can be re-written as

$$X_t = \lambda^2 X_{t-2} + \lambda e_{t-1} + e_t = \cdots = \lambda^t X_0 + \sum_{k=0}^{t-1} \lambda^k e_{t-k}.$$

Gaussian processes – other examples

So for the AR(1) process we have

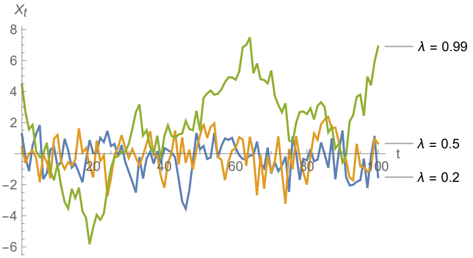
$$
\begin{aligned}
E[X_t] &= E\left[\lambda^t X_0 + \sum_{k=0}^{t-1} \lambda^k e_{t-k}\right] = \lambda^t E[X_0] + \sum_{k=0}^{t-1} \lambda^k E[e_{t-k}] \\
&= \lambda^t E[X_0]
\end{aligned}
$$

and

$$
\begin{aligned}
\operatorname{var}(X_t) &= \operatorname{var}\left(\lambda^t X_0 + \sum_{k=0}^{t-1} \lambda^k e_{t-k}\right) \\
&= \lambda^{2t} \operatorname{var}(X_0) + \sum_{k=0}^{t-1} \lambda^{2k} \operatorname{var}(e_{t-k}) \quad \text{(independence)} \\
&= \lambda^{2t} \operatorname{var}(X_0) + \sigma^2 \sum_{k=0}^{t-1} \lambda^{2k}.
\end{aligned}
$$

Example (AR(1) process).
Below is a Mathematica example for a variety of AR(1) processes.

```
X0 = 0;  σ = 1; TT = 100;
data = Table[RandomFunction[ARProcess[{λ}, σ], {0, TT}], {λ, {0.2, 0.5, 0.99}}];
ListPlot[data, Joined → True, PlotRange → All, AxesLabel → {"t", "Xt"},
 PlotLabels → {"λ = 0.2", "λ = 0.5", "λ = 0.99"}
]
```



Gaussian processes – other examples

Exercise.
Consider the so-called ergodic AR(1) process, that is with $|\lambda| < 1$, and assume $X_0 \sim N(0, \frac{\sigma^2}{1-\lambda^2})$, Show that

$$Q(t,s) = \mathrm{cov}(X_t, X_s) = \frac{\sigma^2}{1-\lambda^2}\lambda^{|t-s|}.$$

Solution. As class work.

Stationary processes

An important class of SPs are those which are stationary, of which there are two definitions.

Definition 8 (strictly stationary SPs)

A stochastic process X_t, $t \in D$, is called "strictly stationary" if for any $n \in \mathbb{N}$

$$F_{X_{t_1},\ldots,X_{t_n}}(x_1,\ldots,x_n) = F_{X_{t_1+h},\ldots,X_{t_n+h}}(x_1,\ldots,x_n)$$

for any $t_i \in D$ and $t_i + h \in D$ with $i \in \{1,\ldots,n\}$.

A weaker form of stationarity exists if the last definition is true for $n \in \{1,2\}$.

Consider case $n = 1$. Then

$$F_{X_t}(x) = F_{X_{t+h}}(x)$$

for any h and so $F_{X_t}(x) = P(X_t \leq x)$ does not depend on t.

Stationary processes

Hence, in particular, for some constant c

$$E[X_t] = E[X_{t+h}] = c.$$

Now consider case $n = 2$. Then

$$F_{X_s, X_t}(x_1, x_2) = F_{X_{s+h}, X_{t+h}}(x_1, x_2) = F_{X_0, X_{t-s}}(x_1, x_2)$$

if $h = -s$.

Then

$$\begin{aligned}
Q(s,t) = \operatorname{cov}(X_s, X_t) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \big((x_1 - E[X_s])(x_2 - E[X_t])\big)\, dF_{X_s, X_t}(x_1, x_2) \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \big((x_1 - c)(x_2 - c)\big)\, dF_{X_0, X_{t-s}}(x_1, x_2) =: q(t-s)
\end{aligned}$$

is a function of $t - s$ only.

These results drive the second definition of stationarity.

Definition 9 (weakly stationary SPs)

A stochastic process X_t, $t \in D$, is called "weakly stationary" if

$$E[X_t] = c \text{ and } \operatorname{cov}(X_t, X_{t+h}) = q(h)$$

where c is a constant.

Obviously, any strictly stationary process is weakly stationary (if, of course, $E[X_t^2] < \infty$).

Example.
Let

$$X_t = \xi \sin(t) + \eta \cos(t), \quad t \geq 0,$$

where ξ and η are uncorrelated RVs with

$$E[\xi] = E[\eta] = 0 \quad \text{and} \quad E[\xi^2] = E[\eta^2] = \sigma^2.$$

Stationary processes

Checking the conditions for weak stationarity we see

$$E[X_t] = E[\xi \sin(t) + \eta \cos(t)] = \sin(t)E[\xi] + \cos(t)E[\eta] = 0$$

and

$$\begin{aligned}
\mathrm{cov}(X_s, X_t) &= \mathrm{cov}\left(\xi \sin(s) + \eta \cos(s), \xi \sin(t) + \eta \cos(t)\right) \\
&= \sin(s)\sin(t)\,\mathrm{cov}(\xi,\xi) + \sin(s)\cos(t)\,\mathrm{cov}(\xi,\eta) \\
&\quad + \cos(s)\sin(t)\,\mathrm{cov}(\eta,\xi) + \cos(s)\cos(t)\,\mathrm{cov}(\eta,\eta) \\
&= \sin(s)\sin(t)\,\mathrm{cov}(\xi,\xi) + 0 + 0 + \cos(s)\cos(t)\,\mathrm{cov}(\eta,\eta) \\
&\text{(independence)} \\
&= \sigma^2(\sin(s)\sin(t) + \cos(s)\cos(t)) \\
&= \begin{cases} \sigma^2 \cos(t-s), & s \le t \\ \sigma^2 \cos(s-t), & t < s \end{cases} \\
&= \sigma^2 \cos(|t-s|).
\end{aligned}$$

Stationary processes

So X_t is weakly stationary (but not necessarily strictly stationary).

However, for Gaussian stochastic processes we have the following result.

Proposition 8 (weakly stationary Gaussian SPs are strictly stationary)

Any weakly stationary Gaussian process X_t is a strictly stationary process.

Proof.
We need only to show that the random vectors

$$\boldsymbol{X} = (X_{t_1}, \ldots, X_{t_n})^T$$

and

$$\boldsymbol{Y} = (X_{t_1+h}, \ldots, X_{t_n+h})^T$$

have the same characteristic function for any $t_i, t_i + h \in D$.

Stationary processes

But the CF of a Gaussian process is completely described by mean and covariance.

Through weak stationarity we know for any $t, t + h \in D$

$$E[X_t] = E[X_{t+h}] = c$$

and due to the property $Q(s, t) = q(t - s)$ for any $s, t, s + h, t + h \in D$

$$\text{cov}(X_s, X_t) = \text{cov}(X_{s+h}, X_{t+h}).$$

Therefore

$$E[\boldsymbol{X}] = E[\boldsymbol{Y}] \text{ and } \text{cov}(\boldsymbol{X}, \boldsymbol{X}) = \text{cov}(\boldsymbol{Y}, \boldsymbol{Y})$$

showing $\boldsymbol{X}$ and $\boldsymbol{Y}$ have the same CF and thereby distribution functions.

References I

Chung, K. (2001).
A Course in Probability Theory.
Academic Press, San Diego, CA, 3rd edition.